Tourism Management 68 (2018) 444–449

Contents lists available at ScienceDirect

Tourism Management

journal homepage: www.elsevier.com/locate/tourman

Opinion Piece

Towards an end of measurement misspecification in tourism research: Grammar of theoretical constructs, focus of thought and mind traps

Josip Mikulić

Faculty of Economics and Business, University of Zagreb, JF Kennedy Square 6, 10000 Zagreb, Croatia

ARTICLE INFO

Keywords:
Construct measurement
Specification
Latent variables
Reflective
Formative
Composite attributes
SEM

ABSTRACT

This article introduces a new perspective on theoretical constructs to help minimizing confusion regarding measurement mode choice in tourism research. In particular, this article uses the grammatical structure of theoretical constructs as contextual framework to explain various mind traps that may mislead proper measurement specification. On the one hand, it is argued that awareness about the precise grammatical structure of constructs may help the researcher to keep the initial focus of thought, which, in turn, helps to avoid reflective *versus* formative misspecification when modeling theoretical constructs. On the other hand, it is further argued that some constructs, in fact, leave the realm of psychological assessments, which is why a direct reflective approach no longer applies. The suggested guidelines are rather simple, but may help clarifying several misconceptions and misunderstandings regarding formative *versus* reflective measurement specification for a range of popular tourism constructs.

1. Introduction

The reflective *versus* formative measurement debate, which is taking place within the area of marketing, management and psychology research during the past decade, has yielded valuable insight that helps researchers obtaining more reliable and valid operationalizations of theoretical constructs (e.g. Baxter, 2009; Becker, Klein, & Wetzels, 2012; Coltman, Devinney, Midgley, & Venaik, 2008; Diamantopoulos, Riefler, & Roth, 2008). To date, several guidelines have been established to assist decision-making about the appropriate measurement mode, involving criteria like e.g. mandatory indicator correlation in reflective mode or flow-of-causality assessments. Works dealing with this issue have appeared in research outlets ranging from *Annals of Tourism Research* (Murphy, Olaru, & Hofacker, 2009) to even *Nature* recently (Kruis & Maris, 2016).

As evidenced by a large number of contemporary tourism studies which rely on latent variable modeling, researchers seem however still to struggle with proper measurement specification. Persistence of the problem is likely associated with the fact that existing guidelines and checklists leave large space for subjective evaluations. For example, whether a manifest indicator is regarded as *causing* a latent variable (formatively identified by the indicator) or as a consequence of the latent variable (reflectively identified by the indicator), remains *a confusing* issue among many tourism researchers. In particular, today there is an abundance of studies that unconsciously force indicators which are obviously formative by nature into a reflective measurement mode. Consequently, results and implications of such studies are highly doubtful while measurement misspecification is further nurtured to appear in future research studies.

The objective of the present article is to clarify this particular issue, which is at the nucleus of measurement misspecifications, and to make a step towards firmer guidelines for the choice of measurement mode in tourism research. To achieve this goal the grammatical structure of theoretical constructs is introduced as contextual framework to discuss and explain various *mind traps* that may mislead proper measurement specification. This article further portrays the situation when tourism research in fact steps outside the realm of classical psychometric theory, which is the major supporting pillar of reflective measurement mode advocates. As this article will show, in such a situation theoretical constructs become inherently formative at second (or highest) level.

The remainder of this article is structured as follows. Section 2 illustrates the grammatical structure of theoretical constructs and portrays situations when the researcher's initial focus of thought may unintentionally shift. This in turn may subsequently lead to reflective *versus* formative confusion. Section 3 portrays the situation when measurements actually leave the realm of classical psychometric approaches and when reflective *versus* formative debates in fact no longer apply. Section 4 provides a set of guidelines for four distinct measurement scenarios which emerge from the discussions. The article concludes with three rather simple but important implications for future research.

E-mail address: jmikulic@efzg.hr.

https://doi.org/10.1016/j.tourman.2018.04.010
Received 28 January 2018; Received in revised form 27 April 2018; Accepted 30 April 2018

Table 1
Grammatical structure of popular theoretical constructs.

Construct name	Subject	Subject's attribute	Preposition	Object
1. Tourist satisfaction	Tourist's	level of satisfaction	with	product/service/brand/destination
2. Destination loyalty	Tourist's	level of loyalty	towards	destination
3. Brand trust	Customer's	level of trust	towards	brand
4. Affective commitment	Customer's	level of affective commitment	towards	product/service/brand
5. Market orientation	Organization's	level of orientation	towards	market
6. Technology readiness	Organization's	level of readiness	towards	(new) technologies
7. Service quality	Service's	level of quality	none	none
8. Supply chain agility	Supply chain's	level of agility	none	none

2. Grammar of theoretical constructs and focus of thought

Table 1 portrays a list of popular theoretical constructs decomposed into their subject, attribute (with preposition) and object, where applicable, which reveals their full spelling and meaning.

This set of illustrative example constructs can be divided into two broad categories. Constructs 1 to 4 are oriented towards customers, while constructs 5 to 8 are oriented to non-living entities. Moreover, the last two examples portray that not every theoretical construct necessarily contains an object.

Following the grammatical structure of theoretical constructs like those in Table 1, is the researcher's focus directed towards the subject, the attribute, or the object? Usually the primary focus is on the attribute. Accordingly, the questions researchers typically aim to answer are how does the *subject's* attribute affect subsequent behaviors or reactions of the subject (i.e. what are the consequences of the attribute), on the one hand, and/or how is the *subject's* attribute being influenced by *something*, on the other hand (i.e. what are the antecedents of the attribute)?

In this regard, the perspective is aligned with the perspective of psychological theory where assessments are focused on psychological attributes of individuals (i.e. of subjects; Cronbach, 1957). Here an attribute of a subject (e.g. an individual) may generally take two distinct forms. In the specific case that a subject's attribute is being conceived as consisting only of a component that is stable regardless of situational effects and/or interactional effects the attribute can be referred to as a *trait* (Steyer, Ferring, & Schmitt, 1992). Conversely, "attributes of individuals that are relatively changeable in nature" could be referred to as *states*. In this latter case, the subject's attribute is a much more dynamic concept that may significantly vary due to and across different stimuli (i.e. due to various situational effects and/or interactional effects). Following these definitions, an attribute of an individual may theoretically take the form of both a trait and a state. Consider the illustrative example of satisfaction as an attribute of a person (i.e. of the subject in our theoretical construct).

2.1. Subject-oriented measurement of attributes

Hypothetically one could think of a person as being generally more or less satisfied regardless of situational and/or interactional effects. In this instance, the attribute (i.e. satisfaction) could be regarded as a trait of the person (i.e. of the subject). If taking a classical, reflective measurement approach an operationalization of this construct should involve measures/indicators that represent observable consequences of satisfaction and, ideally, they should be *object-free*. In case of self-assessments using Likert-type scales such indicators may involve general states and reactions of the subject caused by the trait like e.g. "Generally, I feel fine", "I am rarely sad", "I laugh a lot", or similar.

However, why should one avoid any objects in indicators? The reason is that including objects in indicators may imply introducing some object to the construct, which in turn would *force* measurements into formative mode. Consider the examples of indicators like e.g. "Generally, I am pleased with my partner" or "Generally, I feel fine about my job". Both indicators appear reasonable indicators of a person's general satisfaction and both indicators bear objects in themselves (i.e. partner and job, respectively). Here the object, which has been implicitly introduced to the construct, is a concept that could be referred to as e.g. *life*, whereby job and partner represent relevant aspects of a person's life. However, while these are certainly important aspects of a person's life, these are certainly not the only important aspects. Accordingly, besides the initial requirement to assure validity and measurement reliability for the subject's attribute (i.e. satisfaction) which is in the actual focus, with inclusion of objects to indicators like above one would further have to assure content validity for the implicitly introduced grammatical object (i.e. life), as a precondition for that measurements of the subject's attribute could be reliable and valid. This in turn would require to somehow define *everything* a person (i.e. the subject) can be more or less satisfied with since leaving out an important aspect of life would "… *make the measurement deficient by restricting the domain of the construct*" (Churchill, 1979). Theoretically, covering the complete conceptual domain of *life* in an exhaustive way is possible but is a hard task to fulfill. However, by using indicators like e.g. "Generally, I feel fine about my job" it becomes in fact impossible to assure measurement reliability and content validity for the subject's attribute in a classical test theory manner, as would be indicated by a high Cronbach alpha. The reason is that indicators like these just merge a reflection of the attribute (i.e. I feel fine←satisfaction) with part of the object (i.e. Life←Job) into one indicator. Thus even if content validity of *life* had been assured somehow (i.e. for the object), this would still not imply that validity and reliability for *satisfaction* (i.e. attribute) has been assured. This was however the departing requirement to fulfill *and* in fact the reason for using a reflective measurement approach in the first place! Before taking the discussion further, let us consider an example of object-oriented measurement of a subject's attribute which is far more common in tourism research.

2.2. Object-oriented measurement of attributes

A person may be generally more or less satisfied but also more or less satisfied with e.g. a hotel. The object provides the context for thinking about the subject's attribute thus creating a nomological network (Cronbach & Meehl, 1955). In this instance the focus of thought becomes necessarily object-oriented and the attribute in fact a state, because of the interactional effect with the hotel. In measurement operationalizations, indicators now necessarily have to be object-oriented or otherwise conjunction to the object of the construct (i.e. the context) is lost.

A classical reflective measurement approach would imply assuring reliability and validity at the levels of both the attribute and the object. If the object is however not an abstract concept from the perspective of the subject (who is also the rater in self-assessments), there is in fact no real need to assure content validity for the object (i.e. the object is valid

at face; concrete singular according to Rossiter, 2002). Accordingly, when measuring the construct "hotel satisfaction" in a reflective manner, one would need to assure content validity and measurement reliability for the attribute of the subject (i.e. satisfaction of the customer-hotel guest). In this regard, indicators might resemble those from the previous example of "general satisfaction" but applied to the object of the construct (e.g. "I am pleased with the hotel", "To choose this hotel was a good decision", "I would recommend this hotel to my friends", or similar). To remain reflective, the indicators must however treat the grammatical object of the construct *as a whole in either an explicit or an implicit way*. Otherwise, measurements are again forced into formative mode and a reflective approach does no longer apply. In such a case, the focus of thought can basically shift into two distinct levels of abstraction.

2.2.1. Introducing sub-objects of an object to indicators of a subject's attribute

On the one hand, the focus may shift to different *constituent parts* of the hotel (e.g. lobby, room, restaurant, etc.). Grammatically speaking, what is measured would be the subject's (i.e. customer's) level of the focal attribute (i.e. level of satisfaction) with regard to the sub-object (e.g. with regard to the lobby) of the object (i.e. of the hotel).

Accordingly, with introduction of sub-objects to indicators, reliable and valid reflective measurement of "hotel satisfaction" would require assuring content validity and reliability for the attribute of the subject (i.e. satisfaction of the customer) but also content validity for the object. However, why take the focus to sub-objects if everybody knows what a hotel is? The only situation when this is reasonable is when the actual research focus is on the subject's levels of the focal attribute with regard to the sub-objects of the object (i.e., the researcher's focus is on satisfaction with lobby, room, restaurant, separately). Otherwise, shifting the focus from the object to its sub-objects is completely unnecessary. Besides potentially interesting but actually unintended insight, the only thing one gets is problems to assure content validity for the grammatical object (i.e. did we cover all *parts* of the hotel?). Accordingly, if the initial focus was not on sub-objects researchers should generally avoid taking it there.

2.2.2. Introducing an object's attributes to indicators of a subject's attribute

On the other hand, when measuring hotel satisfaction the focus may also shift to various characteristics of the hotel, like e.g. its design, comfort, location, etc. Grammatically speaking, what is measured would be the subject's (i.e. the customer's) level of the attribute (i.e. level of satisfaction) with regard to the object's level (i.e. with regard to the hotel's level) of some of object's attribute (e.g. of design).

If introducing an object's attributes to indicators, then reliable and valid measurement of "hotel satisfaction" would require assuring validity and reliability for the attribute of the subject (i.e. satisfaction of the customer), content validity for the object (i.e. theoretically all characteristics of the hotel have to be covered) and measurement reliability *and* validity for the attributes of the object (e.g. the design of the hotel). The latter requirement emerges from potentially decreased face validity at the level of an object's attributes (e.g. does everybody think of the same when speaking about design or comfort?). However, if not bothering about face-validity at the level of the object's attributes for a while, then indicators like e.g. "The hotel has an appealing design", or "I am pleased with the hotel's comfort-level", or even "I am satisfied with the hotel's location" really may appear as being appropriate indicators *reflecting the subject's* attribute with regard to the object, which would justify a reflective measurement approach. This is, however, a *mind trap*! What these indicators truly reflect are attributes of the *object* (i.e. the design, comfort-level and location of the hotel), while the attribute of the *subject* (i.e. satisfaction) is only merged with the object's attributes in indicators. In the first two indicators implicitly (i.e. "has an appealing design", "pleased with hotel's comfort-level"), in the latter explicitly ("satisfied with location").

Accordingly, what has happened is that the initial focus of thought has in fact shifted from the attribute's consequences to the attribute's causes (i.e. the object's attributes), which in turn implies a formative measurement approach. If researchers are however not aware that the measurement mode necessarily shifts from reflective to formative when introducing sub-objects or an object's attributes to indicators, then finding three or more correlated measures which are related to sub-objects or attributes of the object might create the misbelief of having achieved reliability and validity for the subject's attribute (e.g. customers are satisfied with their hotel's design, comfort-level and location, across a sample of e.g. ten analyzed customers). This would however only be spurious reliability and validity. The subject's attribute is in fact being completely disregarded from a reflective measurement perspective! Not to say that an indicator like "The hotel has a rich offering of food and meals" might probably be dropped if the accommodation facility was a garni hotel.

3. The construction of meaning

The previous section has illustrated a *mind trap* which is activated when the grammatical object of a construct is not treated as a whole in indicators. In such cases a reflective measurement approach is no longer feasible, but the measurement mode necessarily becomes formative. This section will introduce a more serious, and likely frequent, problem contributing to reflective *versus* formative measurement confusion. For the purpose of illustration another popular construct will be discussed, i.e. *Market orientation*.

3.1. "By-the-book" operationalization

The precise grammatical structure of the construct *Market orientation* would be an organization's (i.e. the subject's) level of orientation (i.e. level of the attribute) towards the market (i.e. the object). Since the object (i.e. market) is rather valid at face, at least for management-oriented researchers, a reflective approach would require to assure content and measurement reliability for the attribute (i.e. "level of orientation" or "orientedness"). Thereby the object needs to be regarded as a whole to avoid shifting into formative mode. Such indicators may for example involve "The organization is much oriented towards the market", "The organization thinks much about the market", "When it is about the market, the organization is highly involved" or similar. These indicators describe possible *reactions* of the subject caused by the focal attribute. Thereby the object is treated as a whole, while an attempt is made to assure validity and reliability for the attribute (i.e. orientedness) by using a reflective measurement approach.

Truly, however, these indicators are all the same indicator but only differently formulated! Compared to the earlier example where researchers may unconsciously arrive at spurious reliability and validity in a psychometric sense, here researchers achieve artificially constructed reliability and validity or what DeVellis (1991) calls *useful redundancy*. More precisely, concurrently valid indicators are artificially created by reformulating the same content. Face-validity of market-orientedness has however certainly not been assured with an operationalization like this. This would however be a necessary precondition for measurement reliability and validity, unless the rater does not exactly know what market-orientedness is or means. Accordingly, a *by-the-book* construct operationalization like this would not be satisfactory though seemingly "reliable" in a psychometric sense.

3.2. Inherent meaning and composite attributes

What is the difference between the two constructs examined so far? The key difference is that the attribute of the construct *Market orientation* (i.e. orientedness) is a concept without inherent, common standalone meaning. *Meaningful, typical, common* direct reflections of the attribute itself cannot be defined. Subsequently, even when in conjunction with a generally meaningful object (i.e. market), full meaning of the whole construct has not been established. This is the reason why it was not possible to define meaningful reflective indicators following the rules set up above (i.e. reflections of the attribute in conjunction with the whole object), except for indicators which in fact contain only reformulated content.

Conversely, the attribute in the construct *Hotel satisfaction* (i.e. satisfaction) has inherent standalone meaning. Reflections of the attribute can be thought of even without the presence of an object. This is because there is a rather common understanding of this attribute. In conjunction with a generally meaningful object (i.e. hotel), meaning of the attribute is maintained (i.e. reinforced), and full meaning of the whole construct *Hotel satisfaction* is established/constructed. Importantly, establishing full meaning of the whole construct is a mandatory condition to fulfill prior to be able to define meaningful consequences, i.e. reflective indicators.

If this is however not feasible like in the example of *orientedness* for the construct *Market orientation*, then the only option is to create a new artificial, *composite attribute* whose meaning researchers would need to *construct* (i.e. define), first. The grammatical structure of the construct could such be collapsed into the following form: an organization's (i.e. a subject's) level of market orientedness (i.e. level of the attribute→ artificial composite attribute), without any grammatical object. Noteworthy, composite attributes can be easily identified because they cannot be brought into conjunction with *any* object (since the object has become part of the artificial attribute).

And at this point researchers might be caught in a *perfidious mind* trap, believing that consequences could be defined prior to defining the meaning/content of the artificial, composite attribute itself. The *idea* that a reflective approach, based on classical test theory, should generally apply is further potentiated by the fact that any first indicator introduced to a construct without an object could in fact be both formative and reflective—i.e., it is no longer possible to specify an indicator as being reflective or formative by looking at how it is related to the object. Consider for example the indicator “The organization cares much about customer needs”. To argue that this indicator well reflects or is a consequence of the subject's attribute (i.e. market-orientedness) would seem reasonable. But consider also the following interpretation. This could be an indicator of an aspect (i.e. an attribute) of market orientedness. This in turn would imply that measurements are located on the antecedents-side of the attribute, while customer care would be one of its formative indicators. Arguing like this would again appear reasonable. If however arguing both the one and the other way appears reasonable, which one of the two approaches would be the more appropriate one?

The answer to this question is that only a formative approach would be appropriate at the highest level of the construct! What is known about the construct *Market orientation* so far is only its name, but not what it actually *is*. And at this point measurements need to step outside the realm of classical psychometric assessments for a while, because the construct must necessarily be modeled formatively, first, in order to *construct its meaning*. Only when this has been done will one (i.e. can one) be able to define meaningful consequences (i.e. reflective indicators). To generalize, constructs built around artificial, composite attributes *are always* formative at second (or their highest) level, like any other construction is formative by definition. Any first indicator which is introduced to the construct may however be regarded as being reflective, but it may thus no longer be part of the construct's conceptual domain (i.e. one of its formative parts). The analogy is not perfect, but proceeding like this when operationalizing constructs would be like starting to build a house from the roof to its foundation. Either way, the construct will be (i.e. must be) formative at the second (or highest) level.

4. Implications and recommendations for construct measurement

Following the above discussions there are four distinct scenarios, which impose particular requirements on the measurement of theoretical constructs:

1. If the attribute of the subject has inherent standalone meaning (i.e. if it is valid at face or implicitly defined to a significant extent) while the object is rather consensually defined, then it is possible to define direct reflections of the construct (e.g. *Tourist's satisfaction with hotel; Customer's trust towards a brand*). In such a situation, reliable and valid construct measurement would require assuring reliability for the attribute by using a reflective measurement mode. A subsequent mandatory requirement for the applicability of a reflective measurement mode would be to treat the object of the construct explicitly or implicitly as a whole in indicators.
2. If the attribute of the subject has inherent standalone meaning while the object is not consensually defined, then it would again be possible to define consequences of the construct (e.g. *Customer's satisfaction with a service*). However, reliable and valid measurement may *primarily* require to assure content validity for the object in a formative mode (i.e. for service), besides assuring reliability for the attribute in a reflective mode (i.e. for satisfaction). This creates a trade-off situation. Generally, however, validity issues should be given higher priority than reliability issues. Ideally, one would divide the construct into two constructs. If using the example of *Service satisfaction*, then satisfaction with the *overall service* may be reflectively identified (i.e. the consequence) while *service performance* (or *service quality*) may be formatively identified (i.e. the cause). Alternatively, researchers may simply explain their survey respondents what they mean by *Service* to achieve consensual agreement upon the object, prior to asking respondents about levels of *Service satisfaction*.
3. If the attribute of the subject does not have inherent standalone meaning, then it is primarily not important to what extent the object is consensually defined or not (e.g. *Organization's level of orientedness towards the market*, *Destination's level of readiness towards smart technologies*, etc.). In such a situation, it is not possible to measure the construct in a direct reflective mode. The attribute and object of the construct should be collapsed into an artificial composite attribute which would be formative by definition at its second (or highest) level (e.g. level of market orientation)! The attribute must no longer be regarded as a *psychological attribute of a person* which could be assessed in a classical test theory manner. Importantly, such an approach does not mean to abandon a critical realist position and taking a classical constructivist position, but only it means that it is not possible to define typical reflections/consequences of something which is not consensually defined/understood, i.e. without inherent standalone meaning. The second (or highest) level in fact becomes the ‘definition level’, which is important to assure content validity of the construct, while it is at the subsequent lower levels where measurement reliability has to be assured (see Fig. 1).
4. If the attribute of the subject does not have inherent standalone

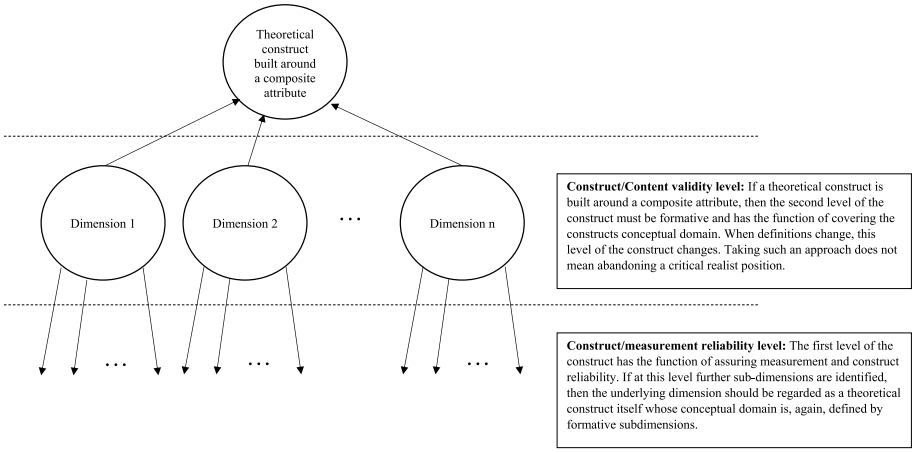

Fig. 1. Structure of constructs built around composite attributes.

meaning while the construct does not contain any object at all, then the same recommendations apply like for the composite attributes above (e.g. *Brand's level of equity*, *Supply chain's level of agility, Service's quality*). A general characteristic of composite attributes is that they cannot be brought into conjunction with any object, which is a sufficient signal to researchers that only a formative measurement mode is applicable at the construct's second (or highest) level.

5. Concluding thoughts

The motivation for writing this article was to help tourism researchers avoiding mind traps that may lead into reflective-formative dilemmas when operationalizing theoretical constructs. Using the grammatical structure of theoretical constructs as a contextual framework to discuss potential problems, this article arrived at three major recommendations.

First, the *grammatical object* of a construct should always be regarded as a whole. If this is not done (i.e. either the focus shifts to sub-objects or sub-attributes), then measurements of the construct necessarily become formative at its second (or highest) level. This in turn leads to potential problems of assuring content validity for the object and may further make measurements overly complicated. If the actual focus does not lie there, it should not be taken there! Moreover, researchers should be aware that finding three or more correlated indicators, which are related to sub-objects and/or an object's attributes, does not justify the use of reflective mode. It is a widespread misconception that this is sufficient for achieving reliability and validity for the subject's attribute. This is however only spurious reliability and validity, because the focus of thought has actually been shifted from the subject's attribute to the construct's object.

Second, if the *grammatical attribute* of a construct does not have inherent standalone meaning, then the construct must (i.e. can only) be modeled formatively at its second (or highest) level. An artificial composite attribute needs to be created whose meaning has first to be "constructed", i.e. defined. Here researchers may step into another mind-trap believing that consequences (i.e. reflective indicators) could be defined prior to establishing full meaning of the composite attribute itself, which is *impossible*. Importantly, if the measurement mode is misspecified (i.e. reflective instead of formative), then the criterion of mandatory indicator correlation will inevitably be violated beyond a certain number of reflective indicators. Below this number however, sufficiently correlated indicators may create the misbelief that construct validity has been achieved (i.e. the Cronbach alpha principle). Above this number, uncorrelated indicators are likely to be discarded in order to obey the reliability imperative. Needless to say that an important formative part of the construct might in fact have been discarded. In analogy to Rossiter's (2005) anecdotal reminder, in which he criticized blind adherence of social science researchers to classical psychometric measurement theory, this article is particularly bothered about the situation when researchers measure e.g. *four legs of the horse* but believe to have captured e.g. the weight (i.e. attribute) of the *whole horse* (i.e. the subject's) very well. High Cronbach alphas would suggest this in a misspecified, reflective measurement mode. Another team of researchers may however focus only on *one leg and* e.g. *the tail* of the horse while both teams would potentially achieve high measurement reliabilities for the composite construct *Horse weight* on a sample of e.g. ten healthy horses. Apparently, this could lead to a situation that virtually any hypothesis could be underpinned with empirical evidence. There is thus a stringent need for both consensual agreement on construct definitions and proper measurement specification in order to be able to create valid theory. The suggested distinction between a content validity level and a subsequent lower reliability level provides simple but effective guidance.

Third and lastly, researchers (and reviewers/editors) should be conscious that the only situation when one has to bother about content validity and measurement reliability for an object or an attribute is when face validity (i.e. full meaning) is not granted. If everyone knows what a hotel is, then content validity does not have to be assured through measurement reliability. This is clear; a *hotel* is not a theoretical construct but a manifest object. However, everyone would probably know what *Satisfaction* means, too. If using reflective approaches to assure measurement reliability for a known attribute, then what is potentially done is decreasing the attribute's content reliability in order to increase measurement reliability, which in turn has the function to increase content validity for the attribute!? Accordingly, this becomes a twisted trade-off situation. In analogy to Rossiter (2005) again, this may potentially lead to the situation that we believe to be measuring *a horse, but what we really obtain is a yak*. If we recall the first example in this article, in which reflective indicators were used for a person's *general satisfaction*, then a closer look at these indicators reveals that they could in fact also underlie a theoretical construct called *Happiness*, rather than *Satisfaction*. Accordingly, unless face validity for the rater is questionable, there is no need for reliability and validity assurance in a classical test theory manner. If face validity is however not granted for the rater (e.g. when using attributes like *Loyalty* or *Commitment*), then researchers should take a classical reflective approach for assuring reliability and validity.

Acknowledgement

This work has been supported by the Croatian Science Foundation under the project DESTBRAND (project no. UIP-2014-09-7005, Period: 2015-2018).

References

Baxter, R. (2009). Reflective and formative metrics of relationship value: A commentary essay. *Journal of Business Research, 62*, 1370–1377.
Becker, J. M., Klein, K., & Wetzels, M. (2012). Hierarchical latent variable models in PLS-SEM: Guidelines for using reflective-formative type models. *Long Range Planning, 45*, 359–394.
Churchill, G. A., Jr. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research, 16*, 64–73.
Coltman, T., Devinney, T. M., Midgley, D. F., & Venaik, S. (2008). Formative versus reflective measurement models: Two applications of formative measurement. *Journal of Business Research, 61*, 1250–1262.
Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist, 12*, 671.
Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52*, 281.
DeVellis, R. F. (1991). *Scale development: Theories and applications.* Newbury Park, CA: Sage.
Diamantopoulos, A., Riefler, P., & Roth, K. P. (2008). Advancing formative measurement models. *Journal of Business Research, 61*, 1203–1218.
Kruis, J., & Maris, G. (2016). Three representations of the Ising model. *Scientific Reports, 6*, 34175.
Murphy, J., Olaru, D., & Hofacker, C. F. (2009). Rigor in tourism research: Formative and reflective constructs. *Annals of Tourism Research, 36*(4), 730–734.
Rossiter, J. R. (2002). The C-OAR-SE procedure for scale development in marketing. *International Journal of Research in Marketing, 19*, 305–335.
Rossiter, J. R. (2005). Reminder: A horse is a horse. *International Journal of Research in Marketing, 22*, 23–25.
Steyer, R., Ferring, D., & Schmitt, M. J. (1992). States and traits in psychological assessment. *European Journal of Psychological Assessment, 8*, 79–98.

Josip Mikulić is associate professor and head of the department of tourism at the Faculty of Economics and Business in Zagreb, Croatia. His research focus is on tourism marketing management and research methods in business and economics.